

10.822.700

07.20.04

日 本 国 特 許 庁
JAPAN PATENT OFFICE

別紙添付の書類に記載されている事項は下記の出願書類に記載されている事項と同一であることを証明する。

This is to certify that the annexed is a true copy of the following application as filed with this Office.

出 願 年 月 日
Date of Application: 2 0 0 4 年 2 月 1 6 日

出 願 番 号
Application Number: 特 願 2 0 0 4 - 0 3 7 5 9 6
[ST. 10/C]: [J P 2 0 0 4 - 0 3 7 5 9 6]

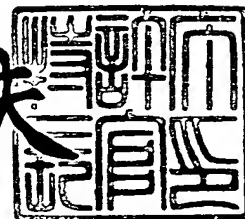
願 人
Applicant(s): 株式会社日立製作所

CERTIFIED COPY OF
PRIORITY DOCUMENT

2 0 0 4 年 4 月 1 4 日

特許庁長官
Commissioner,
Japan Patent Office

今 井 康 夫



【書類名】 特許願
【整理番号】 K03010821A
【あて先】 特許庁長官殿
【国際特許分類】 G06F 12/00
【発明者】
 【住所又は居所】 神奈川県川崎市麻生区王禅寺 1099 番地 株式会社日立製作所
 システム開発研究所内
 【氏名】 松並 直人
【発明者】
 【住所又は居所】 神奈川県川崎市麻生区王禅寺 1099 番地 株式会社日立製作所
 システム開発研究所内
 【氏名】 藺田 浩二
【発明者】
 【住所又は居所】 神奈川県川崎市麻生区王禅寺 1099 番地 株式会社日立製作所
 システム開発研究所内
 【氏名】 山本 彰
【特許出願人】
 【識別番号】 000005108
 【氏名又は名称】 株式会社 日立製作所
【代理人】
 【識別番号】 100075096
 【弁理士】
 【氏名又は名称】 作田 康夫
【選任した代理人】
 【識別番号】 100100310
 【弁理士】
 【氏名又は名称】 井上 学
【手数料の表示】
 【予納台帳番号】 013088
 【納付金額】 21,000円
【提出物件の目録】
 【物件名】 特許請求の範囲 1
 【物件名】 明細書 1
 【物件名】 図面 1
 【物件名】 要約書 1

【書類名】 特許請求の範囲**【請求項 1】**

コンピュータ及び第 2 の記憶装置システムと接続される第 1 の記憶装置システムであって、
第 1 の記憶装置と、
前記第 1 の記憶装置システムを制御する第 1 の制御装置と、
制御装置と制御装置と接続される第 2 の記憶装置とを有する前記第 2 の記憶装置システム
への入出力を制御する、第 2 の制御装置とを有し、
前記第 1 の制御装置は、前記第 2 の記憶装置にファイルシステムを構築することを特徴と
する第 1 の記憶装置システム。

【請求項 2】

請求項 1 記載の第 1 の記憶装置システムであって、
前記第 2 の制御装置は、前記第 2 の記憶装置内に構築された論理ユニットを、前記第 1 の
記憶装置システムが有する RAID グループとして管理し、
前記第 1 の制御装置は、前記 RAID グループを用いて、前記第 2 の記憶装置にファイルシ
ステムを構築することを特徴とする第 1 の記憶装置システム。

【請求項 3】

請求項 1 記載の第 1 の記憶装置システムであって、
前記第 1 の制御装置は、前記コンピュータからファイルへのアクセス要求を受信すると、
前記ファイルが格納されるファイルシステムが、前記第 2 の記憶装置に構築されているこ
とを特定し、
前記第 2 の制御装置を介して、前記第 2 の記憶装置に格納されているファイルへのアクセ
スを実行することを特徴とする第 1 の記憶装置システム。

【請求項 4】

請求項 1 記載の第 1 の記憶装置システムであって、
前記第 2 の制御装置は、複数の第 2 の記憶装置システムと接続されることを特徴とする第
1 の記憶装置システム。

【請求項 5】

請求項 1 記載の第 1 の記憶装置システムであって、
前記第 1 の制御装置は、前記第 1 の記憶装置にファイルシステムを構築し、前記第 1 の記
憶装置に構築されたファイルシステムと前記第 2 の記憶装置に構築されたファイルシステ
ムを、コンピュータから単一のファイルシステムとして認識されるよう管理することを特
徴とする第 1 の記憶装置システム。

【請求項 6】

請求項 1 記載の第 1 の記憶装置システムであって、
前記第 1 の制御装置は、前記第 1 の記憶装置にファイルシステムを構築し、前記第 1 の記
憶装置に構築されたファイルシステムのデータを、前記ファイルシステムの構造を維持して
、前記第 2 の記憶装置内へコピーすることを特徴とする第 1 の記憶装置システム。

【請求項 7】

請求項 6 記載の第 1 の記憶装置システムであって、
前記第 1 の制御装置は、前記第 1 の記憶装置に構築されたファイルシステムのデータを、
前記第 2 の記憶装置内へコピーした後、前記第 1 の記憶装置に構築されていた前記ファ
イルシステムのデータの消去すること特徴とする第 1 の記憶装置システム。

【請求項 8】

請求項 6 記載の第 1 の記憶装置システムであって、
前記第 1 の制御装置または前記第 2 の制御装置は、前記第 1 の記憶装置に構築されたファ
イルシステムのデータを前記第 2 の記憶装置内へコピーする際、前記コンピュータからの
前記ファイルシステムへの書き込みを禁止する旨の属性を付加することを特徴とする第 1
の記憶装置システム。

【請求項 9】

請求項 8 記載の第 1 の記憶装置システムであって、
前記第 1 の記憶装置システムは、管理装置と接続され、前記第 1 の制御装置は、前記管理装置からの指示に基づいて、ファイルシステムの構築と、前記第 1 の記憶装置に構築されたファイルシステムのデータの前記第 2 の記憶装置内へのコピーと、前記属性の付加を、実行することを特徴とする第 1 の記憶装置システム。

【請求項 10】

請求項 9 記載の第 1 の記憶装置システムであって、
前記管理装置により前記第 1 の記憶装置に新たなファイルシステムを構築する時刻が設定されており、
前記時刻になった際、前記第 1 の制御装置は、前記第 1 の記憶装置に新たなファイルシステムを構築し、
前記コンピュータから書き込まれるデータを、前記新たなファイルシステム内に格納し、
前記時刻以前に前記コンピュータから書き込まれるデータを格納していた前記第 1 の記憶装置に構築されたファイルシステムを、前記第 2 の記憶装置内へ移動することを特徴とする第 1 の記憶装置システム。

【請求項 11】

請求項 10 記載の第 1 の記憶装置システムであって、
前記第 1 の記憶装置システムは、管理テーブルを有し、
ファイルシステムの構造を、前記管理テーブルにより管理することを特徴とする記憶装置システム。

【請求項 12】

コンピュータからアクセスされるファイルを記憶するシステムであって、
第 1 の記憶装置システムと第 2 の記憶装置システムを有し、
前記第 1 の記憶装置システムは、前記コンピュータと接続され、
第 1 の記憶装置と、
前記第 1 の記憶装置システムを制御する第 1 の制御装置と、
前記第 2 の記憶装置システムへの入出力を制御する第 2 の制御装置とを有し、
前記第 2 の記憶装置システムは、制御装置と、制御装置と接続される第 2 の記憶装置とを有し、
前記第 1 の制御装置は、前記第 2 の記憶装置にファイルシステムを構築し、
前記コンピュータから前記第 2 の記憶装置システム内に構築されたファイルシステムへのアクセス要求を受け付けて、該ファイルシステムにアクセスすることを特徴とするシステム。

【請求項 13】

請求項 12 記載のシステムであって、
前記第 1 の制御装置は、前記第 1 の記憶装置にファイルシステムを構築し、
前記第 1 の記憶装置に構築されたファイルシステムと前記第 2 の記憶装置に構築されたファイルシステムを、コンピュータから単一のファイルシステムとして認識されるよう管理することを特徴とするシステム。

【請求項 14】

請求項 12 記載のシステムであって、
更に、管理装置を有し、
前記管理装置は、前記第 1 の制御装置に、前記第 1 の記憶装置に構築されたファイルシステムを前記第 2 の記憶装置内へ移動するよう指示し、
前記第 1 の制御装置は、前記指示に従い、前記第 1 の記憶装置に構築されたファイルシステムのデータを、ファイルシステムの構造を維持して、前記第 2 の記憶装置内へコピーし、
前記第 1 の記憶装置システムの記憶装置に構築されたファイルシステムのデータを消去し、
前記コンピュータから前記ファイルシステムへのアクセス要求を受け付けた場合に前記第

1の制御装置は、前記第2の記憶装置内にアクセスするよう制御することを特徴とするシステム。

【請求項15】

請求項14記載のシステムであって、
前記第1の制御装置が、前記第1の記憶装置に構築されたファイルシステムのデータを、
前記第2の記憶装置内へコピーする際、
前記管理端末は、前記第1の記憶装置システムに前記ファイルシステムのデータへの前記コンピュータからの書き込みを禁止する旨の属性を付加するよう指示し、
前記第1の制御装置または第2の制御装置は、
前記ファイルシステムに前記属性を付加することを特徴とするシステム。

【請求項16】

請求項15記載の計算機システムであって、
前記管理装置は、新たなファイルシステムを構築する時刻を設定し、前記時刻になった際、
前記第1の制御装置は、前記第1の記憶装置に新たなファイルシステムを構築し、
前記時刻以降にコンピュータから受信したデータは、前記新たなファイルシステムに書き込み、
前記新たなファイルシステムを構築する以前に前記コンピュータから書き込まれるデータを格納していた前記第1の記憶装置に構築されたファイルシステムを、前記第2の記憶装置内に移動することを特徴とするシステム。

【請求項17】

コンピュータと、第2の記憶装置システムとに接続される、第1の記憶装置システムで実行されるプログラムであって、
第1の記憶装置システムが有する第1の記憶装置への入出力を制御するためのコードと、
第2の記憶装置システムが有する制御装置に対して、前記第2の記憶装置システムが有する第2の記憶装置への入出力を制御するためのコードと、
前記第2の記憶装置にファイルシステムを構築するためのコードを有することを特徴とする第1の記憶装置システムのプログラム。

【請求項18】

請求項17記載の第1の記憶装置システムのプログラムであって、
前記プログラムは更に、
前記第1の記憶装置にファイルシステムを構築するためのコードと、
前記第1の記憶装置に構築されたファイルシステムのデータを、ファイルシステムの構造を維持して、前記第2の記憶装置内にコピーするためのコードを有することを特徴とする第1の記憶装置システムのプログラム。

【請求項19】

請求項18記載の第1の記憶装置システムのプログラムであって、
前記プログラムは更に、
前記第1の記憶装置に構築されたファイルシステムのデータを、前記第2の記憶装置内にコピーした後、
前記第1の記憶装置システムに構築されていた前記ファイルシステムのデータを消去するためのコードを有することを特徴とする第1の記憶装置システムのプログラム。

【請求項20】

請求項18記載の第1の記憶装置システムのプログラムであって、
前記プログラムは更に、
前記第1の記憶装置に構築されたファイルシステムのデータを、前記第2の記憶装置内へコピーする際、前記第1の記憶装置に構築されたファイルシステムへの前記コンピュータからの書き込みを禁止する旨の属性を付加するためのコードを有することを特徴とする第1の記憶装置システムのプログラム。

【書類名】明細書

【発明の名称】記憶装置システム

【技術分野】

【0001】

本発明は、計算機システムで用いられる記憶装置システムに関する。

【背景技術】

【0002】

特許文献1には、階層記憶装置とよぶ計算機に、高速記憶装置と低速記憶装置を接続したシステムが開示されている。特許文献1では、使用頻度の高いファイルは磁気ディスク装置などの高速記憶装置に格納され、使用頻度の低いファイルはテープ装置などの安価な低速記憶装置に格納される。そして、ファイル毎のアクセス頻度を管理するテーブルを用いて、どのファイルをどの記憶装置に配置する、即ち格納するかが決定されている。

【0003】

【特許文献1】特開平9-297699号公報（第3-4頁、図1）

【発明の開示】

【発明が解決しようとする課題】

【0004】

特許文献1では、計算機上で動作するソフトウェアにより、小容量の磁気ディスク装置と、大容量の磁気テープ装置の間でファイルを使用頻度に応じて移動させる階層記憶制御を実現している。階層記憶制御においては、過去にアクセス頻度が高かったデータは、将来もアクセス頻度が高いものと推定し、データのアクセス頻度に関する統計情報や、高速にアクセスできる記憶装置の使用容量に応じて、データを記憶させる記憶装置が決定される。そして、アクセス頻度の高いデータが高速にアクセスできる記憶装置に存在する確率を高めることで、処理効率の向上と大量のファイル管理を実現している。

【0005】

従来の技術の課題は、容量スケーラビリティが計算機に接続できる磁気ディスク装置の台数や磁気テープ装置の台数により制限され、時間と共に増加していく大量の情報を長期に渡り十分に管理できないことである。

【0006】

そこで、大量のファイル情報を長期に渡り保管できる容量スケーラビリティを有する記憶装置システムを開示する。

【課題を解決するための手段】

【0007】

記憶装置システムは、外部記憶装置システムを接続し外部記憶装置システムに対する入出力を制御できるよう構成されており、また、記憶装置システムは、記憶装置システムの内部に備える記憶領域と外部記憶装置システムが備える記憶領域に対してファイルシステムを構築する。

【発明の効果】

【0008】

記憶装置システムは、容量スケーラビリティを有するNASを構築することができる。

【発明を実施するための最良の形態】

【0009】

以下、本発明の実施例について、図面を参照して説明する。尚、本発明は、本実施例に限られるものではない。

【実施例1】

【0010】

まず、第1の実施例について説明する。

(1) システム構成の一例

図1は、第1の実施形態における、計算機システムの一例を示した図である。以下、xは任意の整数を表す。

【0011】

記憶装置1は、ディスクコントローラ11（以下、「DKC11」と呼ぶ）と、複数の磁気ディスク装置17xx（以下、「ディスク17xx」と呼ぶ）と、管理装置18とを有するディスクアレイシステムである。本実施形態において、ディスク17xxは、Fibre Channel（以下、「FC」と呼ぶ）型のインタフェースを備えるFCディスクであるとする。

【0012】

次に、記憶装置1のDKC11の構成の一例を説明する。DKC11は、一又は複数のNASチャネルアダプタ110x（以下、「CHN110x」と呼ぶ）と、一又は複数の異種記憶装置接続アダプタ111x（以下、「CHD111x」と呼ぶ）と、複数のディスクアダプタ12x（以下、「DKA12x」と呼ぶ）と、共有メモリ13（以下、「SM13」と呼ぶ）と、共有メモリコントローラ15（以下、「SMC15」と呼ぶ）と、キャッシュメモリ14（以下、「CM14」と呼ぶ）と、キャッシュメモリコントローラ16（以下、「CMC16」と呼ぶ）とを有する。

【0013】

CHN110xは、ローカルエリアネットワーク20（以下、「LAN20」と呼ぶ）と接続される計算機40x（以下、ホスト40xと呼ぶ）と、ファイルI/Oインタフェースで接続されるインタフェース制御装置である。

【0014】

CHD111xは、ストレージエリアネットワーク30（以下、「SAN30」と呼ぶ）に接続される他の記憶装置50x（以下、「異種記憶装置50x」と呼ぶ）と、ブロックI/Oインタフェースで接続されるインタフェース制御装置である。以下、CHN及びCHDをまとめてチャネルアダプタ（以下、「CH」と呼ぶ）と呼ぶ。

【0015】

SMC15は、CHN110x、CHD111x、DKA12x及びSM13と接続される。SMC15は、CHN110x、CHD111x、及びDKA12xとSM13との間のデータ転送を制御する。

【0016】

CMC16は、CHN110x、CHD111x、DKA12x及びCM14と接続される。CMC16は、CHN110x、CHD111x、及びDKA12xとCM14との間のデータ転送を制御する。

【0017】

SM13は、ボリューム管理テーブル131を有する。ボリューム管理テーブル131は、論理的な一対の連続アドレス空間を備えた内部的な記憶領域の論理構成単位である「論理デバイス（以下、LDEVと呼ぶ）」の構成を記憶し、管理する。

【0018】

ディスク17xxは、DKA12xに接続される。各DKA12xは、DKA12xに接続される一又は複数のディスク17xxへの入出力を制御する。

記憶装置1では、すべてのCHが、CMC16又はSMC15を介して、CM14、SM13、全てのDKA12x、及びすべてのディスク17xへアクセスすることができる。

【0019】

管理装置18は記憶装置1の内部でDKC11と接続されており、各CHや各DKAを介して記憶装置の構成管理を行う。構成情報はSM13に格納され、各CHおよび各DKAが構成情報を共有する。

【0020】

異種記憶装置50xは、記憶装置1の外部に設置された記憶装置1とは異なる種類の記憶装置である。異種記憶装置50xは、SAN30を介してCHD111xに接続される。異種記憶装置50xから見ると記憶装置1がI/Oを発行するホスト計算機の位置づけとなる。

尚、異種記憶装置50xは、以下の説明では記憶装置1とは異なる種類の記憶装置であることとするが、この応用の形態として、異種記憶装置50xは、記憶装置1と同一種の記憶装置であっても良い。

LAN20は、CHN110xとホスト40xとを接続する。一般的に、LANにはIPネットワークが用いられる。

【0021】

SAN30は、CHD111x及び異種記憶装置50xとを接続する。一般的に、SANにはFibre Channel (FC)が用いられる。しかしSANとして、IPネットワークを用い、SANに接続される機器間でSCSIプロトコルに従ったSCSIコマンドをIPパケットでカプセル化して送受信するiSCSIを用いることとしても良い。本実施形態において、SAN30は、異種記憶装置50xを接続するための専用のSANであるものとし、SAN30にホスト40xは接続されていないものとする。

【0022】

管理端末600は、管理LAN70を介して、記憶装置1が備える管理装置18と接続される。また、管理端末600は、管理LAN70を介して、異種記憶装置50xとも接続される。管理端末600では、記憶装置1や異種記憶装置50xの設定、管理を行うための管理ソフトが動作する。

【0023】

尚、図1に示す記憶装置1は、ホスト40xとLAN経由で接続するためのNASインタフェース (CHN110x)のみを有するが、本実施形態は、記憶装置1がSAN経由でホストに接続するためのSANインタフェース (SANチャネルアダプタ)も備えて、NASインタフェースとSANインタフェースの双方を選択可能な構成とすることも可能である。

(2) 記憶装置の外観構成の一例

図2は、記憶装置1の外観の一例を示した図である。

【0024】

DKCユニット19は、DKC11の構成要素である、CHN110x、CHD111x、DKA12x、SM13、及びCM14を格納する。SM13は、実際には複数のコントローラボード13xで構成される。また、CM14も、複数のキャッシュボード14xで構成される。記憶装置1の使用者は、これらのボードの枚数を増減して、所望の記憶容量のCM14又はSM13を有する記憶装置1を構成することができる。ディスクユニット(以下、「DKU」と呼ぶ)180及びDKU181は、複数のディスク17xxを格納する。

【0025】

スロット190には、CHN110x、CHD111x、DKA12x、コントローラボード13x、キャッシュボード14x等が作りこまれたアダプタボードが格納される。本実施形態においては、スロット190の形状、アダプタボードのサイズ及びコネクタの形状を、アダプタボードの種類やインタフェースの種類を問わず一定にし、互換性を保つようにする。したがって、DKCユニット19には、アダプタボードの種類、インタフェースの種類を問わず、任意のスロット190に任意のアダプタボードを装填することができる。また、記憶装置1の使用者は、CHN110x及びCHD111xのアダプタボードの数を自由に選択して、DKCユニット19のスロットに選択した数のCHN110x及びCHD111xを装填することができる。

(3) CHNボードの外観構成の一例

図3は、CHN110xが作りこまれたアダプタボード(以下、CHNボードと呼ぶ)の外観構成の一例を示した図である。コネクタ11007は、DKCユニット19が有するコネクタと接続される。インタフェースコネクタ2001は、LAN20と接続可能である。

【0026】

本実施形態においては、アダプタボードのコネクタの形状は、アダプタボードの種類によらず一定であるため、CHNボードとCHD111xが作りこまれたアダプタボード(以下、CHDボードと呼ぶ)とは同一形状のコネクタを有する。尚、CHDボードの場合、インタフェースコネクタ2001は、ファイバチャネルに対応し、ファイバチャネルと接続できるように構成されている。

(4) NASチャネルアダプタ (CHN) の構成の一例

図4は、CHN110xの構成の一例を示した図である。ファイルアクセス制御用CPU11001は、ファイルアクセスを制御するプロセッサである。LANコントローラ11002は、インタフェースコネクタ2001を介してLAN20と接続され、LAN20との間のデータの送受信を制御する。ファイルアクセス制御用メモリ11004は、ファイルアクセス制御用CPU11001と接続される。ファイルアクセス制御用メモリ11004には、ファイルアクセス制御用CPU11001が実行するプログラムや、制御データが格納されている。

【0027】

ディスクアレイ制御用CPU11008は、ディスクアレイを制御するプロセッサである。ここで、ディスクアレイとは、複数台のディスク17xxでグループを構成し、このグループを一台の仮想的なディスクとして扱い、複数のディスク17xxを並列動作させることで性能を高めることができる構成である。特に、このグループの記憶領域の一部にパリティと呼ぶ冗長データを格納することで耐障害性を高めたディスクアレイのことをRAID (Redundant Arrays of Inexpensive Disks) と呼び、ディスクアレイの中でも特によく用いられている。

【0028】

ディスクアレイ制御用メモリ11009は、ディスクアレイ制御用CPU11008と接続され、ディスクアレイ制御用CPU11009が実行するプログラムや、制御データが格納されている。SM I/F制御回路11005は、CHN110xからSM13へのアクセスを制御する回路である。CM I/F制御回路11006は、CHN110xからCM14へのアクセスを制御する回路である。CPU間通信回路11007は、ファイルアクセス制御用CPU11001がディスクアクセスのためにディスクアレイ制御用CPU11008と通信する際に用いられる通信回路である。

【0029】

なお、本実施形態では、CHN110xにファイルアクセス制御用CPU11001とディスクアレイ制御用CPU11008の2つのプロセッサが搭載された非対象型のマルチプロセッサ構成の例を示しているが、ファイルアクセス制御とディスクアレイ制御を単一のプロセッサで実行するようCHNを構成して実装することもできるし、ファイルアクセス制御とディスクアレイ制御の一または両方を二つ以上のプロセッサで対等に実行する対称型マルチプロセッサ構成としてCHN110xを実装することもできる。

(5) ファイルアクセス制御用メモリの構成の一例

図5は、CHN110xが有するファイルアクセス制御メモリ11004に格納されているプログラムや制御データの一例を示した図である。オペレーティングシステムプログラム110040は、プログラム全体の管理や入出力制御に用いられる。LANコントローラドライバプログラム110041は、LANコントローラ11002の制御に用いられる。TCP/IPプログラム110042は、LAN上の通信プロトコルであるTCP/IPの制御に用いられる。ネットワークファイルシステムプログラム110044は、記憶装置に格納されるファイルを、NASホスト40xに提供するためのプロトコルであるNFSやCIFS等の制御に用いられる。ボリューム制御プログラム110045は、1または複数の論理ユニット（以下、「LU」と呼ぶ）から論理ボリュームを構成する制御に用いられる。CPU間通信ドライバプログラム110046は、ファイルアクセス制御用CPU11001とディスクアレイ制御用CPU11008間で通信を行うためのCPU間通信回路11007の制御に用いられる。

【0030】

ファイルシステムプログラム110043は、記憶装置に格納されるファイルの管理に用いられ、ファイルの格納管理や入出力制御を実行するプログラムであり、具体的には、

- 1) ファイルを使用する際にファイルをオープンする。
 - 2) ホストからファイルアクセス要求を受信した際にアクセス要求に従ったディスクへの入出力処理を実行する。
 - 3) ファイルのディスク上での格納位置を決定し管理する。
 - 4) オープンされたファイルのファイル名と、そのファイルのファイル格納位置やバッファアドレスを管理するテーブルとの対応関係を管理する。
- といった制御処理を行うプログラムである。

(6) ディスクアレイ制御用メモリの構成の一例

図6は、CHN110xが有するディスクアレイ制御メモリ11009に格納されているプログラムや制御データの一例を示した図である。

【0031】

オペレーティングシステムプログラム110090は、プログラム全体の管理や入出力制御に用いられる。

【0032】

CPU間通信ドライバプログラム110093は、ファイルアクセス制御用CPU11001とディスクアレイ制御用CPU11008との間で通信を行うためのCPU間通信回路11007の制御に用いられ、ファイルアクセス制御用CPU11001からのLUに対するアクセス要求を受信する。

【0033】

ボリューム制御プログラム110092は、複数のディスク17xxによりRAIDを構成したRAIDグループ（以下、「VDEV」と呼ぶ）上に、論理的な一対の連続アドレス空間を備える記憶領域の論理構成単位である論理デバイス（以下、「LDEV」と呼ぶ）を一つまたは複数構成し、このLDEVを一つまたは複数結合して論理ユニット（以下、「LU」と呼ぶ）を構成し、その関係情報を管理する。

【0034】

キャッシュ制御プログラム110094は、CM14に格納されているデータの管理や、キャッシュヒット／ミス判定等の制御に用いられる。

【0035】

DKA通信ドライバプログラム110095は、ディスク17xxへのアクセスが必要となった際に、DKA12xと通信するために用いられる。

【0036】

ディスクアレイ制御プログラム110091は、CPU間通信ドライバプログラム110093を介して受信したファイルアクセス制御用CPU11001からのLUに対するアクセス要求から、ボリューム制御プログラム110092によりファイルアクセス制御用CPUからアクセスされるLUに対応する、LDEV、VDEVを特定し、キャッシュ制御プログラム110094により当該アクセスのキャッシュヒットミス判定を行い、ディスクアクセスが必要な際にDKA通信ドライバプログラム110095によりDKA12xにアクセス要求を発行するという一連のディスクアレイ制御を行う。

(7) 異種記憶装置接続アダプタ(CHD)の構成の一例

図7は、CHD111xの構成の一例を示した図である。

【0037】

異種記憶装置接続制御用CPU11108は、異種記憶装置50xを接続するための制御を行うプロセッサである。

【0038】

異種記憶装置接続制御用メモリ11109は、異種記憶装置接続制御用CPU11108と接続され、異種記憶装置接続用CPU11108が実行するプログラムや、制御データが格納されている。SM I/F制御回路11105は、CHD110xからSM13へのアクセスを制御する回路である。CM I/F制御回路11106は、CHD110xからCM14へのアクセスを制御する回路である。

(8) 異種記憶装置接続制御用メモリの構成の一例

図8は、CHD111xが有する異種記憶装置接続制御用メモリ11109に格納されているプログラムや制御データの一例を示す図である。

【0039】

オペレーティングシステムプログラム111090は、プログラム全体の管理や入出力制御に用いられる。

【0040】

異種記憶装置接続制御プログラム111091は、SAN30に接続されている異種記憶装置50xを認識し、異種記憶装置50xが提供するLUの容量を確認したり、LUに対してリード処理やライト処理を行う制御プログラムである。

【0041】

ボリューム制御プログラム111092は、異種記憶装置50xが提供するLUを記憶装置1が備えるVDEVの一つとしてみなし管理する制御プログラムである。このLUは記憶装置1内部ではVDEVとして扱われるので、このVDEV上にLDEVが構成される。ボリューム制御プログラム111092で異種記憶装置50x内のLUがVDEVとして管理されると、このVDEVは、CHN110xのボリューム制御プログラム110092でLDEVと対応づけられ、更にこのLDEVは、LUと対応付けられる。

なお、異種記憶装置50x内のLUはVDEVとして扱うが、異種記憶装置50xがRAID型ディスクアレイ装置の場合、その異種記憶装置50x内でRAIDが構成されているため、記憶装置1が冗長データを付加する必要はない。

【0042】

キャッシュ制御プログラム111093は、CM14に格納されているデータの管理や、キャッシュヒット／ミス判定等の制御に用いられる。

【0043】

LDEVマイグレーション制御プログラム111094は、LDEVのマイグレーションを行う制御プログラムである。

【0044】

記憶装置1が管理するLDEVのデータの中身を異種記憶装置50x上に構成したLDEVに複写し、元のデータを消去すると、あたかもLDEVが移動したように扱うことができる。このような処理をLDEVマイグレーションと呼ぶ。LDEVマイグレーション制御プログラム111094は、記憶装置1と異種記憶装置50x間のLDEVマイグレーションや、異種記憶装置50x内部のLDEVマイグレーションや、一台の異種記憶装置50xと別の一台の異種記憶装置50x間のLDEVマイグレーションを実行する。

【0045】

WORM制御プログラム111095は、CHD1110が管理する異種記憶装置50xに構成したLDEVに対しWORM (Write Once Read Many) 属性を付与する制御を行うプログラムである。WORM制御の一例としては、マイグレーションをした際のライト処理以外のライト処理はすべて抑止し、リードオンリーLDEVとして扱うことでWORM属性を実現する制御が考えられる。WORM制御としては、その他にも、LDEVに追記のみ許可することでWORM属性を実現する制御もある。以下、本発明ではマイグレーションをした際のライト処理以外のライト処理はすべて抑止し、リードオンリーLDEVとして扱う方法を前提に説明する。

(9) 異種記憶装置の構成の一例

図9は、異種記憶装置50xの構成の一例を示した図である。異種記憶装置50xは、ディスクアレイコントローラFCTLx510xを有している。耐障害性を備えるためにFCTLを2枚備えており、一方が障害で停止した際には他方が肩代わりするよう構成されている。

【0046】

ディスクアレイコントローラFCTLx510xは、FCコントローラ51012と、ディスクアレイ制御用CPU51008と、ディスクアレイ制御用メモリ51009と、キャッシュメモリ51014と、データ転送制御回路51011と、FCコントローラ51010を有する。FCコントローラ51012は、FCTLx510xをSAN30に接続される。

ディスクアレイ制御用CPU51008はディスクアレイを制御するプロセッサである。ディスクアレイ制御用メモリ51009には、ディスクアレイ制御用プログラムや制御データが格納されている。FCコントローラ51010はディスク5710xを接続し制御する。データ転送制御回路51011は、FCコントローラ51012とCM51014とFCコントローラ51010の間に配置され、CM51014へのデータの入出力や、他方のFCTLへのデータ転送を制御する。

【0047】

異種記憶装置50xは、記憶装置1と同様に、複数のディスク5710xによりRAIDグループを構成し、その一部もしくは全部の記憶領域に一連のアドレスを配した1または複数のLUを構成する。

【0048】

図9において、記憶装置1のCHD0(1110)は、異種記憶装置0(500)に異種記憶装置接続SAN30を介して接続されており、異種記憶装置0(500)から見ると、記憶装置1のCHD0(1110)はあたかも1台のホストとして動作する。CHD0(1110)は、異種記憶装置0(500)のホストとしてリードやライトといったアクセスコマンドを発行し、異種記憶装置0(500)を制御する。

(10) 異種記憶装置接続SANに1台の異種記憶装置を接続した構成の一例

以上の構成に基づき、以下、本実施形態の動作を説明する。

【0049】

図10は、上記で説明した構成の詳細を省き、特徴だけを記載した計算機システムの簡略構成の一例を示した図である。

【0050】

図11は、図10に対応するボリューム管理テーブル131の状態の一例を示した図である。

【0051】

図11(a)に示すボリューム管理テーブル131は、異種記憶装置500で構築されるLUと、異種記憶装置で構築されるLUから構成されるVDEVの関係の一例を示した図である。図11(a)に示すボリューム管理テーブル131において、SLUNは、異種記憶装置500内で構築されるLUの異種記憶装置500内における識別番号、即ち、異種記憶装置500内のディスクアレイコントローラFCTLx510xがこのLUにアクセスする際に用いる識別番号を示す。VDEVは、異種記憶装置500内で構築されるLUに対応するVDEVを示す。

【0052】

図11(b)に示すボリューム管理テーブル131は、図11(a)で示すVDEVと、LDEVの関係の一例を示した図である。図11(b)に示すボリューム管理テーブル131において、LDEVは、VDEVに対応するLDEVを示す。

図11(c)に示すボリューム管理テーブル131は、図11(b)で示すLDEVと、LUN、LV、FSの関係の一例を示した図である。図11(c)に示すボリューム管理テーブル131において、LUNは、LDEVに対応するLUの番号、LVは、LUに対応する番号、FSは、LVに対応するFSを示す。

【0053】

尚、本実施例では、ボリューム管理テーブル131がFSの情報を持ち、管理するとしたが、FSは、別の管理テーブルで管理してもよく、本実施例に限られるわけではない。

【0054】

図10の記憶装置1において、図4の示すCHN1110の構成を総してNAS機能1100Aと呼ぶ。また、図7に示すCHD1110の構成を総して異種記憶装置接続制御機能1110Aと呼ぶ。

【0055】

論理デバイス5200(SLDEV0)が、異種記憶装置500のSLUN=0のLUに構築され、ファイルシステム5300(SFS0)が、この論理デバイスSLDEV0上に構築される場合を例に本実施形態の動作を説明する。

【0056】

尚、異種記憶装置500は、異種記憶装置500のディスクアレイコントローラFCTLx510xにより、SLUN=0のLUを構築しているとする。

【0057】

まず、図8を参照してCHD0(1110)の異種記憶装置接続制御機能1110Aの動作を説明する。異種記憶装置接続制御プログラム111091は異種記憶装置500にInquiryコマンドを発行することで、異種記憶装置500内のLU0を検出する。ボリューム制御プログラム111092は、このLU0をSVDEV0とみなしSM13のボリューム管理テーブル131に登録する。即ち、ボリューム制御プログラム111092は、図11(a)に示すボリューム管理テーブル131のSLUNに0を、VDEVにSVDEV0を登録する。

【0058】

次に、図6を参照してCHN0(1100)のNAS機能1100Aの動作を説明する。図6のディスクアレイ制御CPU11008が実行するボリューム制御プログラム110092はSM13に格納されている図11(a)に示すボリューム管理テーブル131を参照し、SVDEV0を検出する。ボリューム制御プログラム110092は、このSVDEV0に適当なサイズのSLDEV0を作成し、SM13に格納されている図11(b)に示すボリューム管理テーブル131に登録する。即ち、ボリューム制御プログラム110092は、図11(b)に示すボリューム管理テーブルのVDEVにSVDEV0を、LDEVにSVDEV0に対応するSLDEV0を登録する。

【0059】

図1の管理端末18から管理者は、所望の容量のLUを構成するための指示を計算機システ

ムに発行し、指示を受信した記憶装置1は、1又は複数のLDEVからLUを構成しておく。ここでは、1つのSLDEV0で1つのLU0を作成したとする。即ち、ボリューム制御プログラム110092は、図11(c)に示すボリューム管理テーブル131のLDEVにSLDEV0を、LUNにSLDEV0に対応するLU0の番号0を登録する。

【0060】

次に、記憶装置1がホスト40xにより起動されると、記憶装置1内の図5のファイルアクセス制御用CPU11001が実行するボリューム制御プログラム110045は、CPU間通信ドライバプログラム110046を用いてディスクアレイ制御CPU11008にInquiryコマンドを発行し、ディスクアレイ制御CPU11008にLU0を検出するための問い合わせを行う。ディスクアレイ制御CPU11008が実行するボリューム制御プログラム110092は、LU0を検出し、ファイルアクセス制御用CPU11001に通知する。ボリューム制御プログラム110045はこのLU0を認識し、LU0を用いて論理ボリュームLV0を構築する。尚、論理ボリュームは複数のLUを結合して構築することができるが、ここでは、一つのLU0でLV0を構築したとする。即ち、ボリューム制御プログラム110045は、図11(c)に示すボリューム管理テーブル131のLVにLU0に対応するLV0を登録する。

【0061】

ファイルシステムプログラム110043は管理者からの指示により、論理ボリュームLV0上にファイルシステムSFS0を構築する。即ち、図11(c)に示すボリューム管理テーブル131のFSにLV0に対応するSFS0を登録する。

【0062】

以上の動作により、本実施例では、記憶装置1のLV0は、記憶装置1のLU0から構築され、LU0はSLDEV0から構築される。そして、SLDEV0は、異種記憶装置システム内のLUから構成される記憶領域SVDEV0から構成されている。この結果、異種記憶装置500上にCHN0(1100)が制御する論理デバイスSLDEV0が構築され、そのSLDEV0上にCHN0(1100)が制御するファイルシステムSFS0が構築される。

【0063】

ホスト0(400)がファイルシステムへの問い合わせを発行すると、図5に示すLANコントローラドライバプログラム110041はLANコントローラ11002を制御してLAN20からこの問い合わせを含むパケットを受領し、TCP/IPプログラム110042およびネットワークファイルシステムプログラム110044の働きにより、この問い合わせをファイルシステムプログラム110043が認識する。そしてファイルシステムプログラム110043はSFS0のディレクトリ情報をホスト0(400)に送信し、ホスト0(400)は図10に示す通りこのファイルシステムを記憶装置1内に存在するSFS0と認識し、以後使用することができるようになる。ここで、SFS0のディレクトリ情報とは、一般的なディレクトリ情報であるため、本明細書に詳細は記載しない。

【0064】

本実施例では、記憶装置1が備えるディスク17xxは一切必要とせず、NAS機能1100Aが作成したファイルシステムSFS0の管理情報は、CHD0の異種記憶装置接続制御機能1110Aの働きにより、異種記憶装置接続SAN30を通り、異種記憶装置50のLU上に構築されたSLDEV0に格納されることになる。ここで、ファイルシステムSFS0の管理情報とは、メタデータと一般によばれる一般的なファイルシステムの管理情報であり、その構造については、本明細書に詳細は記載しない。

【0065】

以上、図10の構成によれば、異種記憶装置接続制御機能1110Aにより記憶装置1に接続した異種記憶装置500が備えるディスク上に、記憶装置1のNAS機能1100Aが、ファイルシステムを構築することができる。このような機能のことを、「異種記憶装置接続NAS機能」と呼ぶことにする。

【0066】

ホスト0(400)が、ファイルシステムFS0に格納されているファイルにアクセスする場合、ホスト0(400)がCHN0(1100)にファイルシステム名とファイル名を指定したアクセス

要求を送る。ファイルシステム名とファイル名を受信したCHN0(1100)は、ボリューム管理テーブル131を参照して、当該ファイルが格納されている場所を確定し、当該ファイルが異種記憶装置に格納されている場合は、CHD0(1110)に異種記憶装置内のデータにアクセスが来ている旨を通知すると同時にボリューム管理テーブル131からえられたデータの格納場所を通知し、CHD0(1110)が異種記憶装置内のデータにアクセスする。また、CHN0(1100)が異種記憶装置内のデータにホストからアクセス要求があるという情報をSM13に格納しておき、CHD0(1110)は、定期的に異種記憶装置内のデータにホストからアクセス要求がきているかどうかの情報を確認し、アクセス要求が来ていることを認識すると、ボリューム管理テーブル131を参照して、CHD0(1110)が、異種記憶装置内のデータにアクセスしてもよい。尚、CHD0(1110)が異種記憶装置500にアクセスする時は、異種記憶装置500のディスクアレイコントローラFCTLx510xが認識可能なアドレスを使用することが必要である。従って、CHD0(1110)は、データが格納されているVDEVに対応するSLUN(ボリューム管理テーブル131(a))を用いてアクセスする。

(11) 異種記憶装置接続SANに複数の異種記憶装置を接続した構成の一例

次に、異種記憶装置接続SANに複数の異種記憶装置を接続した構成の一例について説明する。

【0067】

図12は、異種記憶装置接続SANに複数の異種記憶装置を接続した構成の一例を示す図である。図10の構成との相違点は、異種記憶装置接続SAN30に複数(同図では2台)の異種記憶装置500、501が接続していることである。

【0068】

図10と同様の動作を、異種記憶装置500および501に対して行うことにより、異種記憶装置500にはSLDEV0上にSFS0が、異種記憶装置501にはSLDEV1上にSFS1が構築され、CHN0のNAS機能1100Aの働きにより、ホスト0(400)からはSFS0およびSFS1が認識される。

【0069】

このように異種記憶装置接続SAN30に接続する異種記憶装置50xの台数を増加していくことで、CHN0が扱うことができる記憶容量を、増加させていくことができる。

【0070】

以上、図12の構成によれば、記憶装置1の異種記憶装置接続SAN30に、任意の台数の異種記憶装置を接続可能となり、このことにより、容量スケーラビリティを有する大規模なNASを構築することができる。

【0071】

以上、本実施形態によれば、異種記憶装置接続NAS機能を構成することができ、また、外部記憶装置により記憶容量を増大させることができるので、大容量のファイルデータを格納できる容量スケーラビリティに優れた大規模NASを構築することができる。また、ホストは、外部記憶装置システムの記憶領域に構築されたファイルシステムであることを意識することなく、外部記憶装置を接続し、外部記憶装置に対する入出力を制御する記憶装置システムを介して外部記憶装置システムのファイルシステムにアクセスできる。

【実施例2】

【0072】

次に、第2の実施例について説明する。

【0073】

本実施形態は、実施例1で説明した、異種記憶装置接続NASの機能を応用した実施形態である。

【0074】

図13は、第一の応用例を示した図である。図13の図10との相違点は、記憶装置1のCHN0が、異種記憶装置500に論理デバイスSLDEV0を構築し、SLDEV0上にファイルシステムSFS0を構築している他、更に記憶装置1のディスク17xx上に論理デバイスPLDEV0(180)を構築し、論理デバイスPLDEV0(180)上にファイルシステムPFS0(190)を構築している点である。

【0075】

また、PFS0のマウントポイントm1に、SFS0のルートディレクトリがマウントされている点も異なる。CHN0は、PFS0のマウントポイントm1に、SFS0のルートがマウントされたことを認識している。ホスト0(400)からは、PFS0とSFS0は一つのディレクトリツリーとして構成された単一のファイルシステムPFS0として認識されるようになる。

【0076】

ここで、PFS0の一部としてSFS0を結合し、あたかもひとつのディレクトリツリーのように見せかける方法のことを、ここではマウントと表すことにする。一例として、SFS0のルートをPFS0のm1のポイントからソフトリンクすることで、SFS0をPFS0に結合することができる。別の方法として、m1にSFS0のルートをマップするようにファイルシステムを構築してもよい。なお、m1のことをマウントポイントと表すことにする。

【0077】

CHN0は、記憶装置1内のVDEVに構築されたあるFSに、記憶装置1内の他のVDEVに構築された他のFSをマウントするのと同様の方法でPFS0のマウントポイントm1にSFS0をマウントすることができる。CHN0は、記憶装置1内のVDEVも異種記憶装置500内のLUに構築されたVDEVも同じVDEVとして扱い、その上にLDEV、LV、FSを構築するからである。

このように、本実施形態では、記憶装置1の内部の論理デバイスPLDEVx上に構築したファイルシステムPFSxに異種記憶装置500に構築した論理デバイスSLDEVxの上に構築したSFSxを組み合わせることで内部のディスク容量だけでは構築できない大容量の単一ディレクトリツリーを備えるファイルシステムPFSxを構築できることを示している。この構成により、ホストは、記憶装置1の内部のファイルシステムであるとか異種記憶装置500上のファイルシステムであるとかといった意識をすることなく、透過的なSingle Viewで大規模なファイルシステムを使用することができるようになる。

また、ホスト0(400)は、PFS0もしくはPFS0にマウントされているSFS0に格納されるファイルにアクセス要求が発行するとき、ファイルシステム名PFS0とファイル名を指定する。ホスト0(400)からのアクセス要求を、CHN0が受け付けると、CHN0は、当該ファイルがPFS0かSFS0のいずれに格納されているかを確認し、SFS0に格納されているなら、以下、実施例1と同様の方法で、異種記憶装置システム内に格納されたファイルのデータにアクセスする。

【実施例3】

【0078】

次に、第3の実施例について説明する。

【0079】

図14は実施例1の第二の応用例を示した図である。図14は基本的に図13と同一であるが、PLDEVおよびSLDEVの数が増えた実施形態である。

【0080】

データの運用の開始年月が2003年1月であるとする、最初の状態では、記憶装置1の内部にPLDEV0、PLDEV1が構築されており、それぞれの論理デバイスの上にファイルシステムPFS0、PFS1が構築されている。PFS0をホームファイルシステムとして、マウントポイントm1にPFS1がマウントされ、Single ViewのファイルシステムPFS0を構成し、ホスト0(400)にファイルサービスを提供している。

【0081】

図15(a)が2003年1月の時点でのボリューム管理テーブル131の状態の一例を示した図である。ボリューム管理テーブル131において、LVはCHN0が認識している論理ボリュームを示す。LUNはCHN0が認識しているLUの番号を示す。Storageは、当該LUが格納されている記憶装置を示し、STR0は記憶装置1を、STR1は異種記憶装置500を示す。LDEVは記憶装置1内の当該LUが構築されているLDEVを示す。本実施例では、すべて単一のLDEVでLUが構築されているとしている。尚、複数のLDEVを結合し一つのLUを構築することもできる。WORMはWrite Once Read Manyの略で、書き込みは一度だけで、読み出しのみ多回数可能な属性を示し、1がWORM属性を有する記憶領域である旨を示す。Migrationは、当該論理ボリューム

がマイグレーションされたか否かを示し、1がマイグレーションされた記憶領域である旨を示す。FSは当該LVに構築されているファイルシステム名を示す。Remarksは補足事項、例えば、ホームファイルシステムであることを示すhome、ファイルシステムに格納されているデータがいつのものを示す日付等を示す。

【0082】

例えば、図15(a)の一行目は、LV0はLUN0により構成され、LUN0は記憶装置1の内部に構築されたPLDEV0から構成され、ファイルシステムPFS0が構成されており（この状態を、LU0/STRO/PLDEV0/PFS0と表現する）、ホームファイルシステムとされていることを示している。同様に図15(a)二行目は、LV1はLU1/STRO/PLDEV1/PFS1という構成を有し、2003年1月のファイルが格納されていることを示している。以下この状態を表現するのに、LV1/LU1/STRO/PLDEV1/PFS1という表記する。

【0083】

ここで、本実施例においては、各LVには同一年月のファイルのみを格納するという例を示す。

【0084】

LDEVマイグレーションの動作を説明する。ここでは、マイグレーションは「月が変わったら先月のファイルを記憶装置1から異種記憶装置にマイグレーションするとともに、WORM属性に変更する」という運用指針で実行する例で説明する。このような運用指針の管理は記憶装置1の外部に設けた管理端末600のメモリに格納されたソフトウェアであるDLCM6001が行う。DLCMはData Life Cycle Managerの略である。DLCM6001が管理端末600によって実行されると、DLCM6001は、管理装置18から管理情報、特にボリューム管理テーブル131の情報を記憶装置1から入手し、記憶装置1の構成がどのような構成になっているかを把握することができる。

【0085】

ここで、例えば、2003年2月1日になったとすると、管理端末600上のDLCM6001は、まず2月に作成されるファイルを格納するためのファイルシステムPFS2を作成するため、管理装置18を介して記憶装置1のCHN0にPLDEV2を作成する旨の実行指示を発行する。この指示をCHN0が受信し、LV2/LUN2/STRO/PLDEV2/PFS2を作成し、ファイルサービスを開始する。

【0086】

次にDLCM6001は、運用指針に基づき、先月「2003年1月」に作成したファイルにWORM属性を設定し、これらのファイルを記憶装置1から異種記憶装置500にマイグレーションすることを決定する。

【0087】

具体的には、DLCM6001は、「2003年1月」に作成したファイルは、図15(a)に示すボリューム管理テーブル131の情報に基づき、LV1/LUN1/STRO/PLDEV1/PFS1に格納されていることを特定する。

【0088】

次にDLCM6001は、管理装置18を介して記憶装置1のDKC11にLV1にWORM属性を設定するための実行指示を発行する。この指示をCHD0のWORM属性プログラム111095が受信し、まず管理装置18からWORM属性を設定するよう指示を受けたLV1に対応するPLDEV1にWORM属性を設定し、新規書き込みを抑止する。尚、本実施形態では、WORM属性プログラム111095が新規書き込みを抑止するとしたが、本実施形態に限られるものではなく、CHN0やCHD0に新規書き込みを抑止する機能を設けてもよい。

【0089】

次にDLCM6001は、マイグレーション先のLDEVとして異種記憶装置500のSLDEV1に決定する。もし、事前にLDEVが存在していないならば、管理装置18を介して記憶装置1のDKC11にLDEV作成指示を発行する。この指示をCHD0が受信し、異種記憶装置接続制御プログラム11101はSLDEV1を作成する。

【0090】

次にDLCM6001は、管理装置18を介して記憶装置1のDKC11にPLDEV1内のデータをSLDEV1に

マイグレーションするための実行指示を発行する。この指示をCHD0が受信し、LDEVマイグレーション制御プログラム11104は、PLDEV1のデータ内容をSLDEV1すなわち異種記憶装置500に対して異種記憶装置接続SAN30を介してコピーする。LDEVマイグレーション制御プログラム11104は、コピーが完了したら、PLDEV1のデータは消去する。マイグレーションが完了すると、CHD0のボリューム制御プログラム111093は、ボリューム管理テーブル131を変更する。

【0091】

ここで、CHN0は、新規にファイルシステムSFS1を構築するのとは異なり、PFS1の情報がそのままSLDEVに移動されたので、必要に応じて行われるファイルシステム名の更新や、その他必要な情報のみ更新することでSFS1を構築することができる。

【0092】

以上で、1月に作成されたファイルシステムPFS1が格納されているLV1のマイグレーションが完了したことになる。この結果、ボリューム管理テーブル131は、図15(b)の状態になる。LV1は、LV1/LUN1/STR1/SLDEV1/SFS1となりWORM属性が設定された状態になっている。

【0093】

以下、同様に、2003年3月になると、DLCM6001は、3月のファイルを格納するためのPFS3を格納するLV3を作成し、また、2月のファイルシステムPFS2が格納されているLV2を記憶装置1から異種記憶装置500にマイグレーションする。その結果、LV2は図15(c)の通り、LV2/LUN2/STR1/SLDEV2/SFS2となりWORM属性が設定された状態になる。以上の結果として、最終的に図14の構成図のような構成になる。

【0094】

このようにPLDEV1はSLDEV1に、PLDEV2はSLDEV2にマイグレーションしたが、この際に、LV1/LU1、LV2/LU2は何ら変更されず、マイグレーション後のファイルシステムは、homeであるファイルシステムPFS0にマウントされているため、ホストからはマイグレーションが実行されたことは認識されない。また、LDEVマイグレーション制御プログラムは、マイグレーション経過中のアドレスポインタを管理しているので、マイグレーション中にホストからアクセスがあっても、アクセスが遮られたり、誤ったデータが転送されるということはない。その結果、ホストからはファイルシステム全体としての完全なる透過性とSingle Viewが確保される。このような機能のことを、「異種ストレージ接続NASによるSingleViewマイグレーション」と呼ぶことにする。

【0095】

上記の説明では、ボリューム作成処理やマイグレーション処理は、管理端末600上のDLCM6001が実行するとして説明したが、DLCM6001の機能をホストのアプリケーションに組み込んだり、ホスト上でDLCM6001を動作させたり、または、記憶装置の内部の管理装置18上でDLCM6001を実行させることもできる。特にホストのアプリケーションにDLCMの機能を組み込んだ場合には、よりアプリケーションの特性や動作に合致した緻密な制御を実行することが可能になる。この場合、DLCM6001と記憶装置1の間のインタフェースはAPIとして規定され、上記で説明した個々の操作がAPIとして提供されることになる。

【0096】

また、上記の説明では、記憶装置1と異種記憶装置500のコントローラの構成以外の相違点について、説明しなかったが、例えば、記憶装置1はFCインタフェースを備えた高性能なディスク17xxを使用し、異種記憶装置500はATAインタフェースを備えた低価格なディスク5710xを使用することで、ファイルシステムを構築してから書き換えが頻発する期間には高性能FCディスクを、書き換えが発生せず、保管が主体となる時期には低コストATAディスクを利用するといった使い分けを行うことができ、目的に添ったコストパフォーマンスの優れた記憶装置システムを構築することが可能である。特に低コストな、ATAディスクを用いて、時間とともにデータをアーカイビングしていく用途の記憶装置を構築することに対して有効であり、ファイルシステムのバックアップ、Emailのアーカイビング、ログのアーカイビング、監視映像のアーカイビング等、各種アプリケーションに適応可能で

あり利用価値は非常に高い。

【0097】

本実施形態により、記憶装置の複数の論理ボリュームに独立したファイルシステムを構築し、これを一つのツリーとして結合構成することでホストにSingle Viewのファイルシステムを提供することができる。

【0098】

また、異種記憶装置上の記憶領域に論理ボリュームを構築し、異種記憶装置接続機能のマイグレーション機能により、ファイルシステムごとSingle Viewを維持したままマイグレーションすることができる。

【0099】

また、マイグレーション時に、論理ボリュームに対しWORM属性を付加し、書き換えを抑制することができる。

【0100】

また、記憶装置の外部の管理端末もしくはホストからの制御により、あらかじめ定めた指針に従ってボリューム作成およびマイグレーションを実行することができる。

【図面の簡単な説明】

【0101】

【図1】 本発明が適用される計算機システムの構成例を示す図である。

【図2】 記憶装置の外観の一例を示す図である。

【図3】 アダプタボードの外観の一例を示す図である。

【図4】 N A S チャンネルアダプタの構成の一例を示す図である。

【図5】 ファイルシステム制御用メモリに格納されているプログラムの一例を示す図である。

【図6】 ディスクアレイ制御用メモリに格納されているプログラムの一例を示す図である。

【図7】 異種記憶装置接続制御用チャンネルアダプタの構成の一例を示す図である。

【図8】 異種記憶装置接続制御用メモリに格納されているプログラムの一例を示す図である。

【図9】 異種記憶装置の構成例を示す図である。

【図10】 記憶装置に一台の異種記憶装置が接続された構成の一例を示す図である。

【図11】 ボリューム管理テーブルの構成の一例を示す図である。

【図12】 複数台の異種記憶装置を接続する構成の一例を示す図である。

【図13】 記憶装置と異種記憶装置を混在して利用する構成の一例を示す図である。

【図14】 記憶装置と異種記憶装置を混在して利用する構成の一例を示す図である。

【図15】 ボリューム管理テーブルの構成の一例を示す図である。

【符号の説明】

【0102】

1…記憶装置

1100…N A S チャンネルアダプタ

1110…異種記憶装置接続制御アダプタ

120…ディスク制御アダプタ

13…共有メモリ

14…キャッシュメモリ

15…共有メモリコントローラ

16…キャッシュメモリコントローラ

1700…ディスク

400…ホスト

500…異種記憶装置

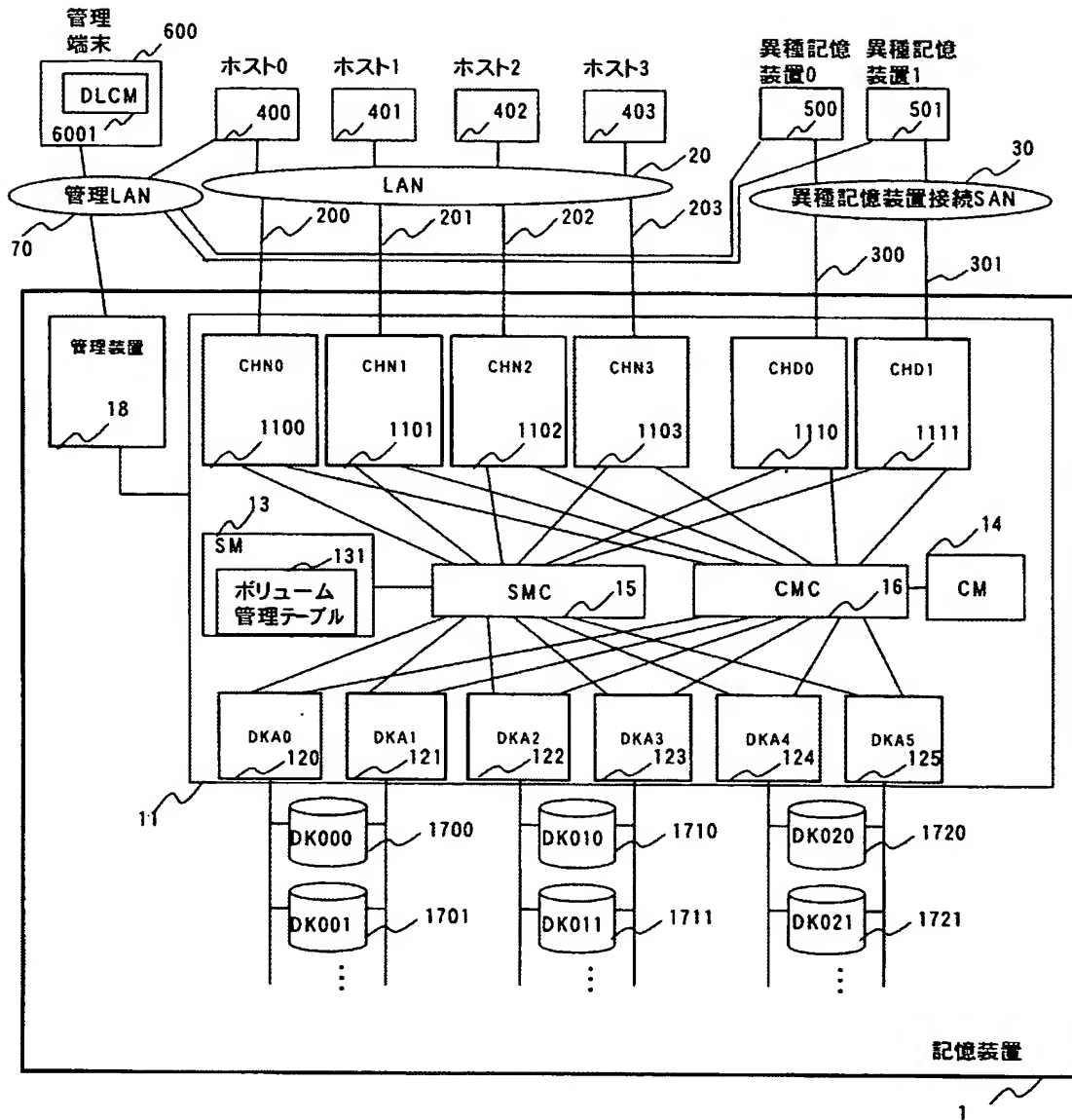
20…LAN

30…異種記憶装置接続SAN

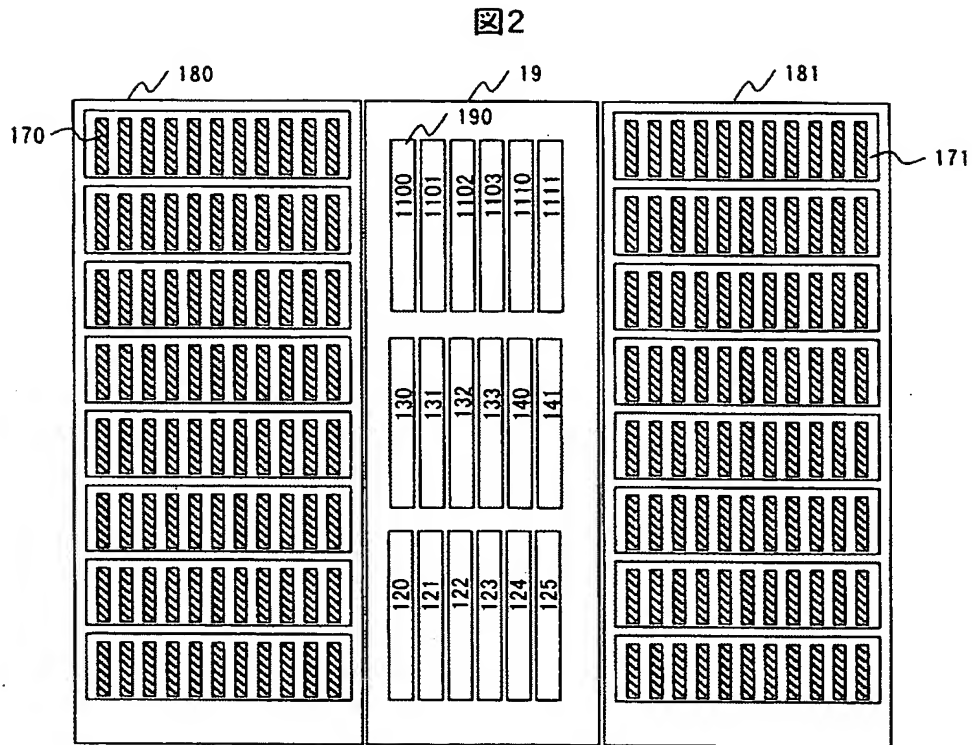
600...管理端末

【書類名】 図面
【図 1】

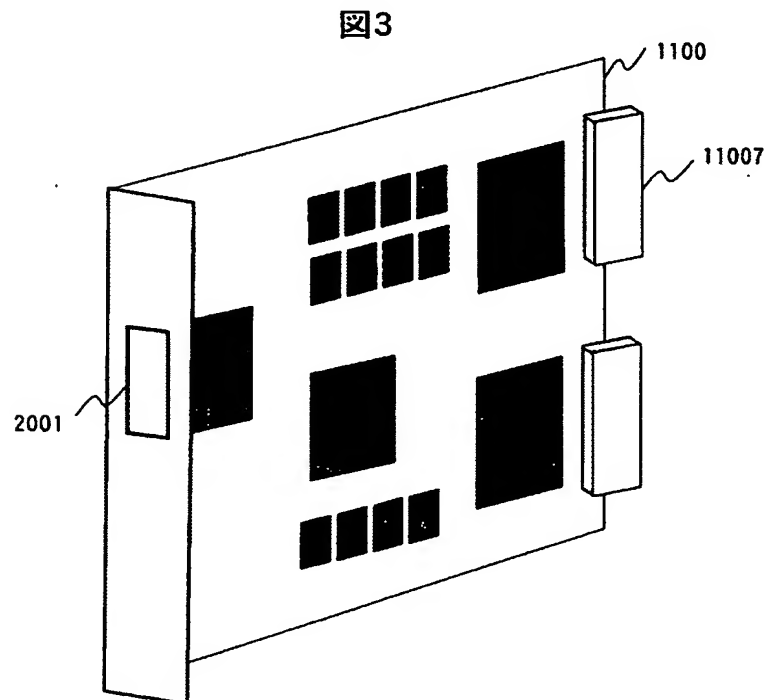
図 1



【図 2】

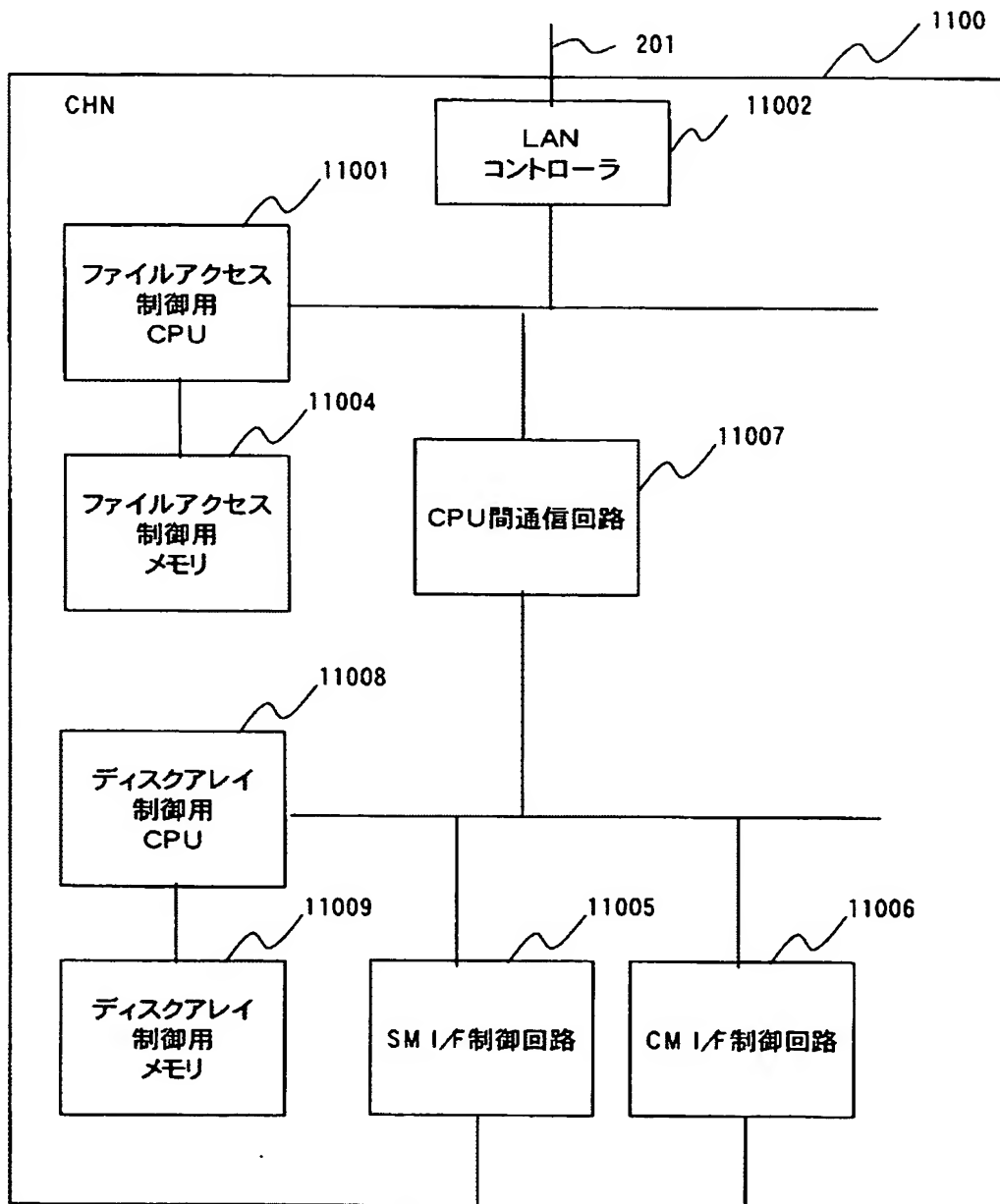


【図 3】



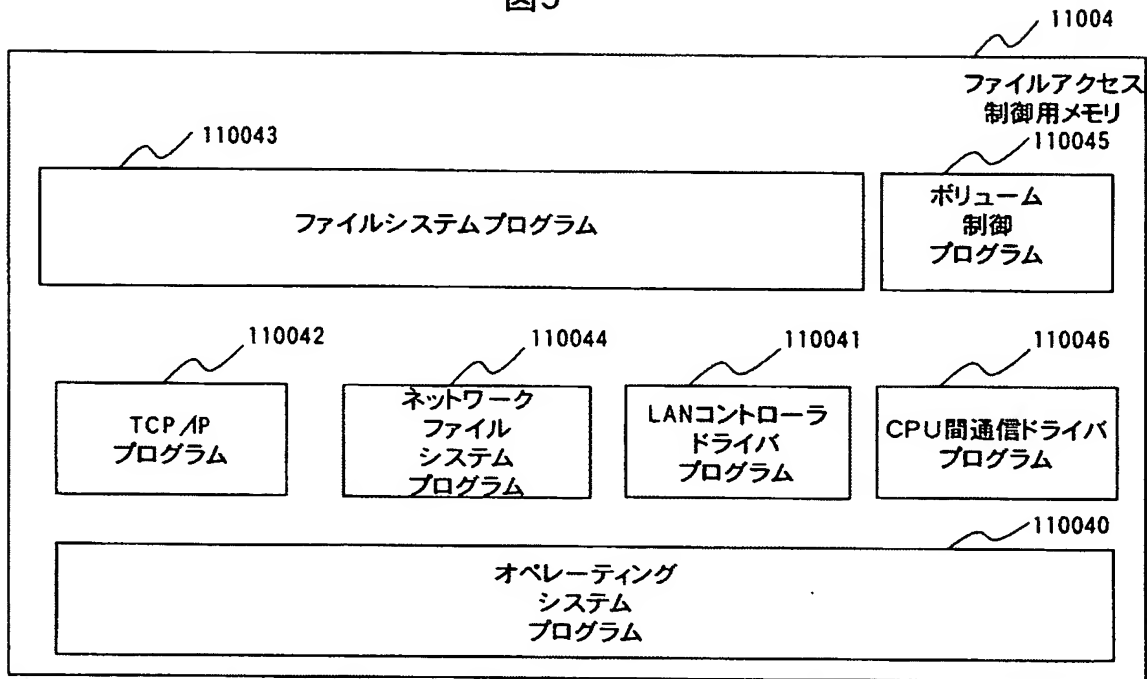
【図 4】

図 4



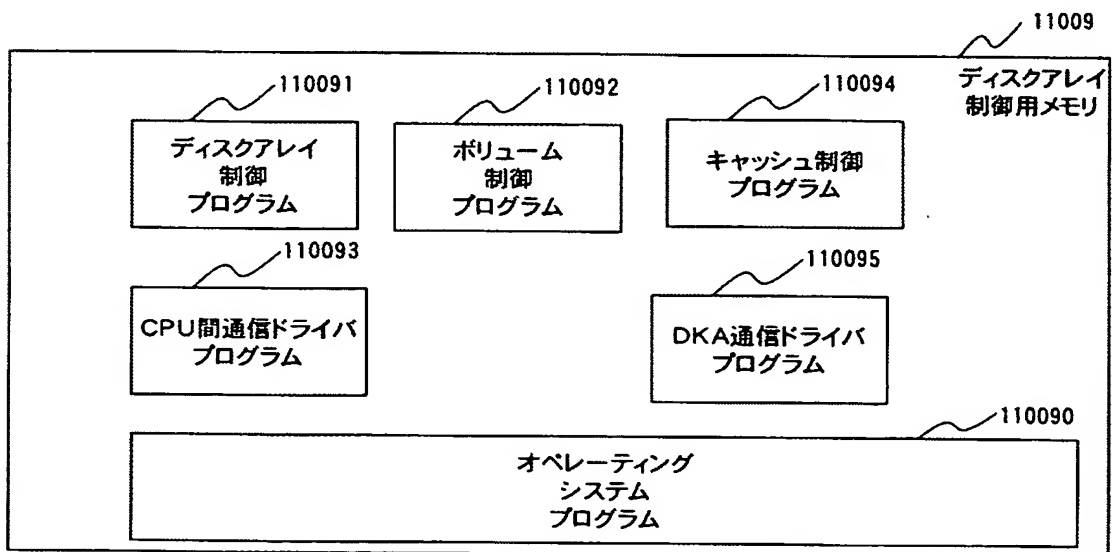
【図 5】

図5

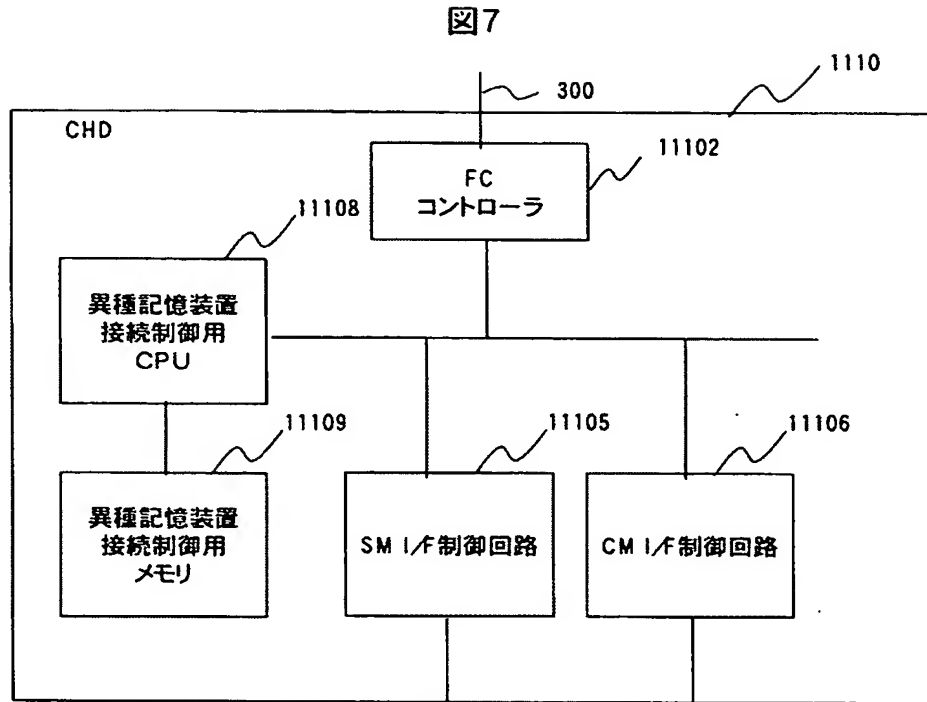


【図 6】

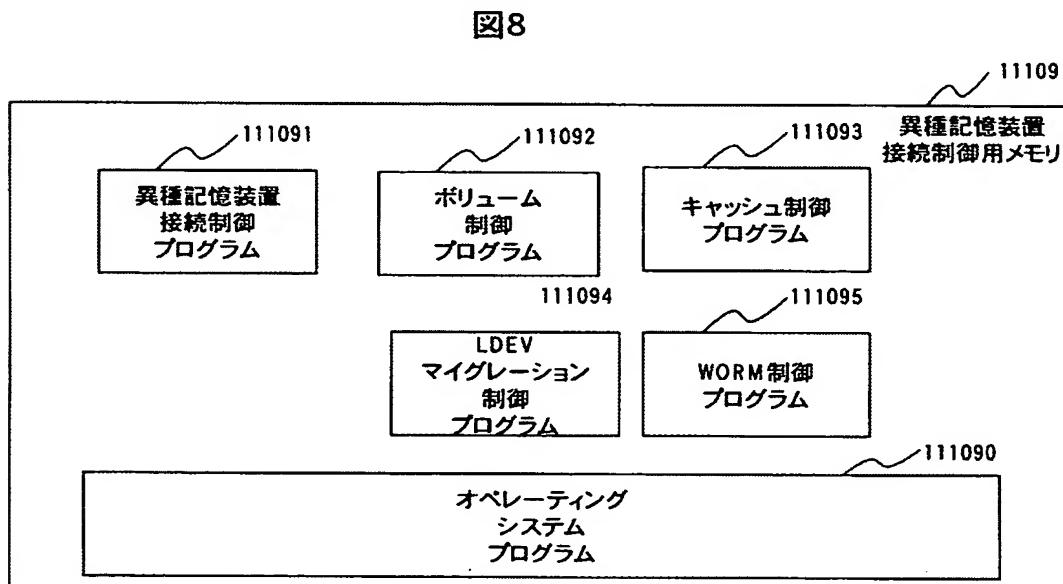
図6



【図 7】

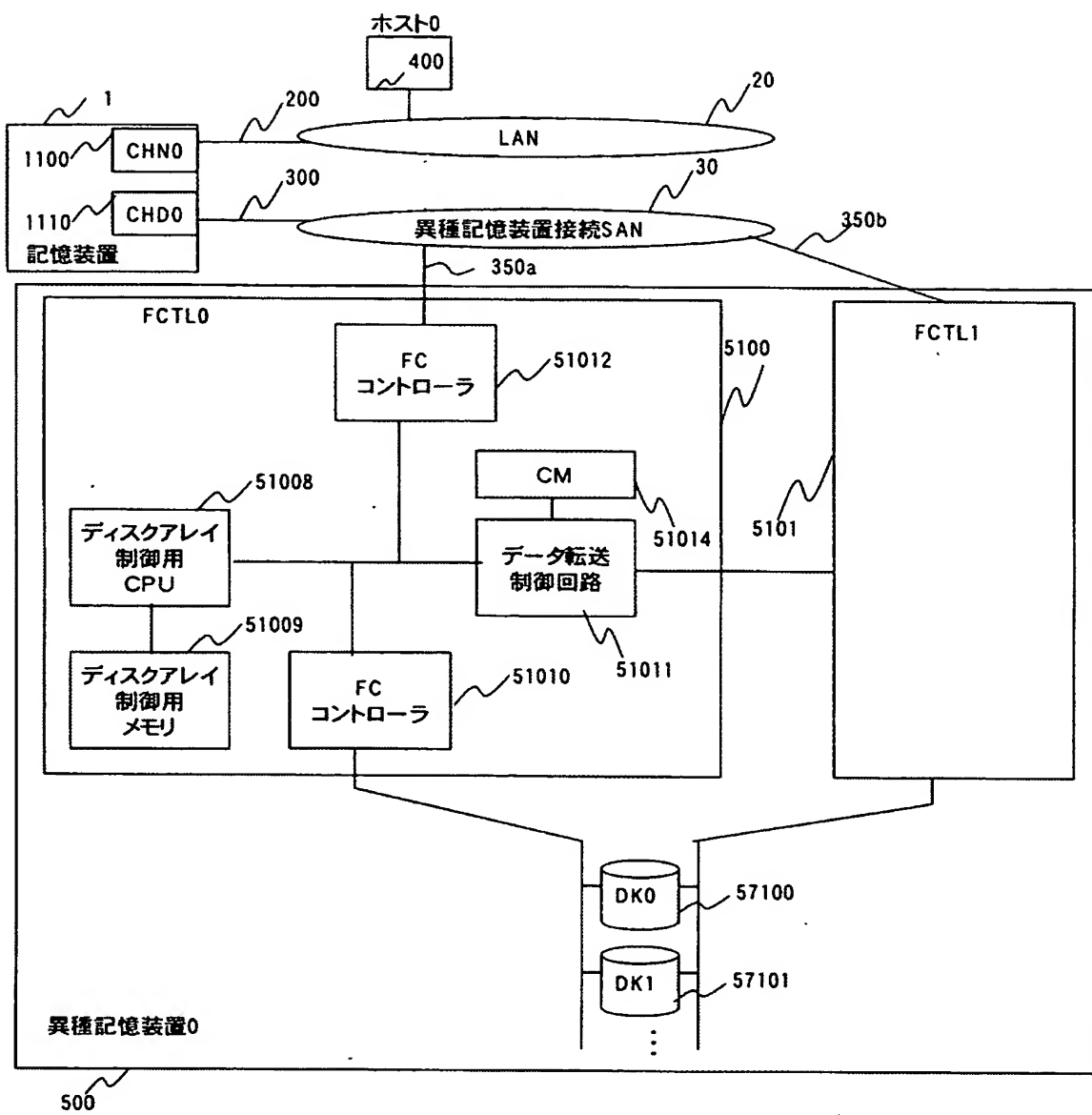


【図 8】



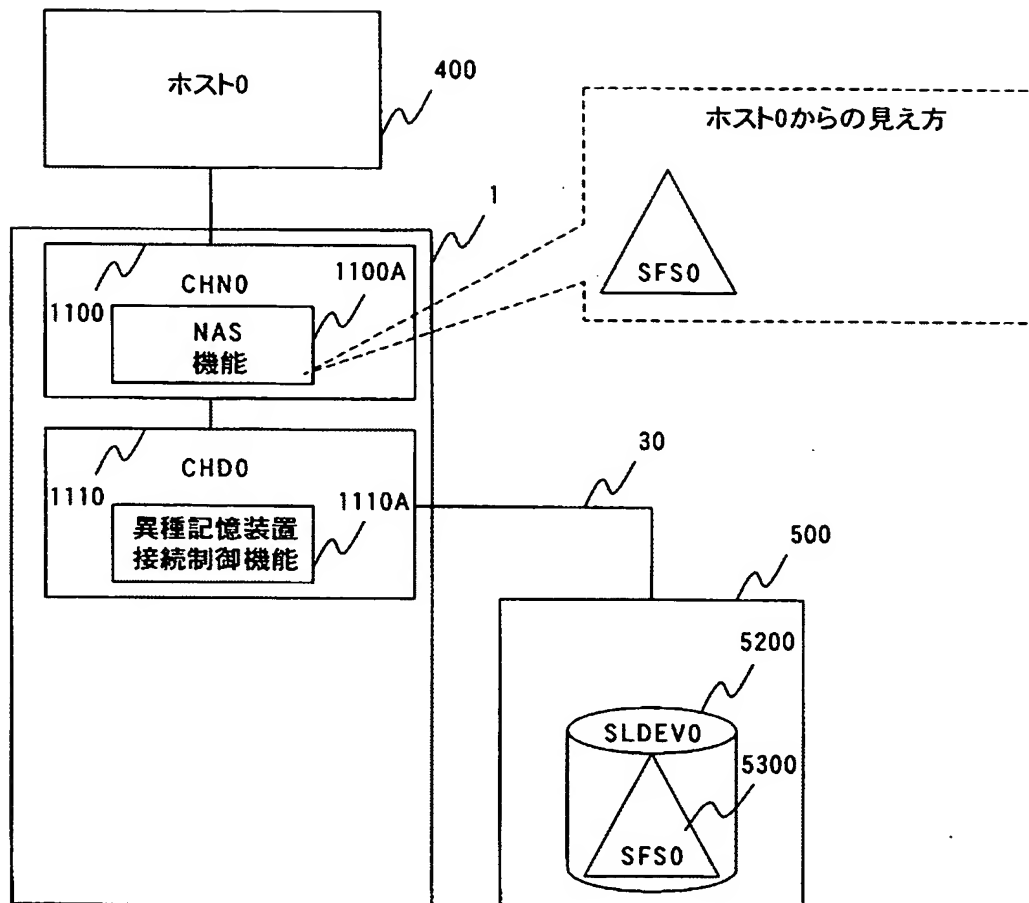
【図 9】

図9



【図 10】

図10



【図 11】

図 11

131

Storage	SLU	VDEV
STR1	0	SVDEV1
:	:	:
:	:	:

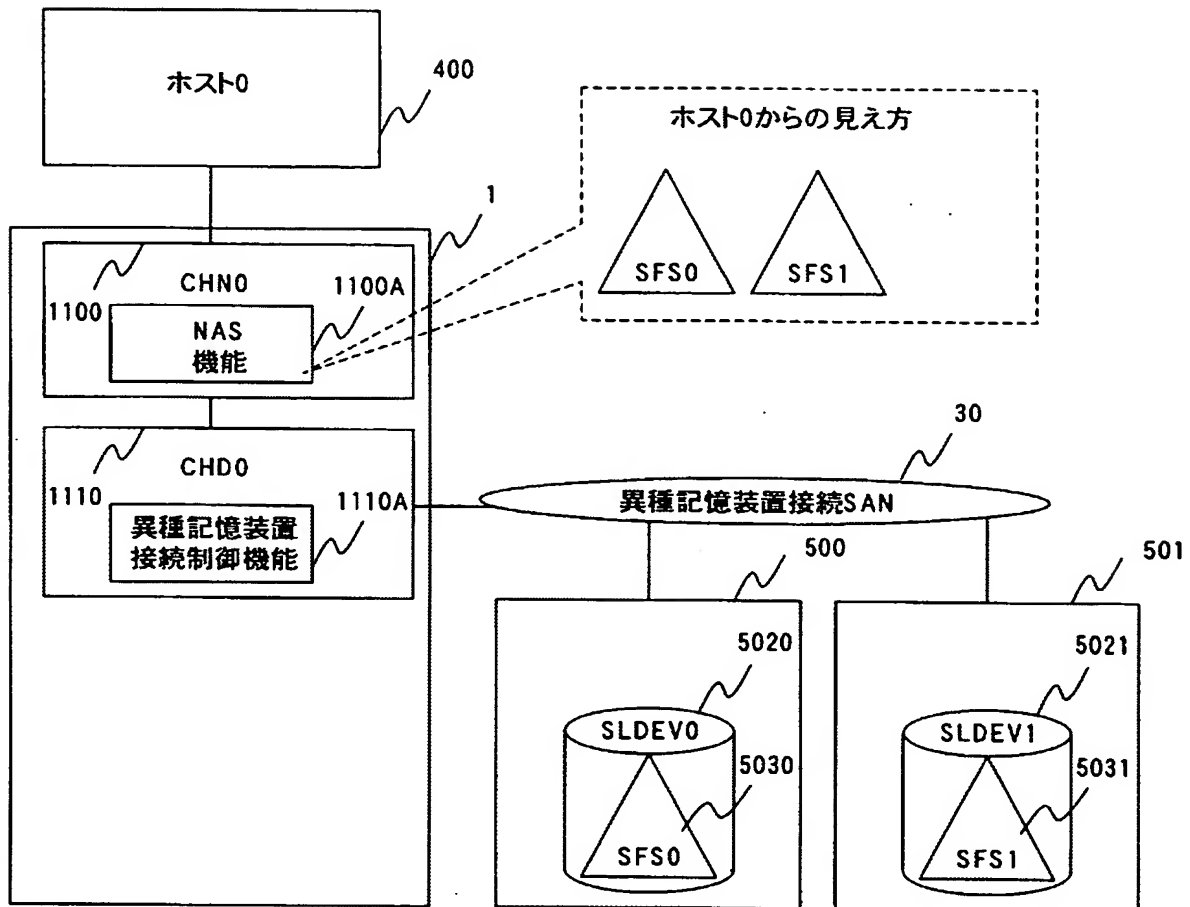
VDEV	LDEV
SVDEV1	SLDEV1
:	:
:	:

LU	LUN	Storage	LDEV	FS
0	0	STR1	SLDEV1	SFS0
:	:	:	:	:
:	:	:	:	:

ボリューム管理テーブル

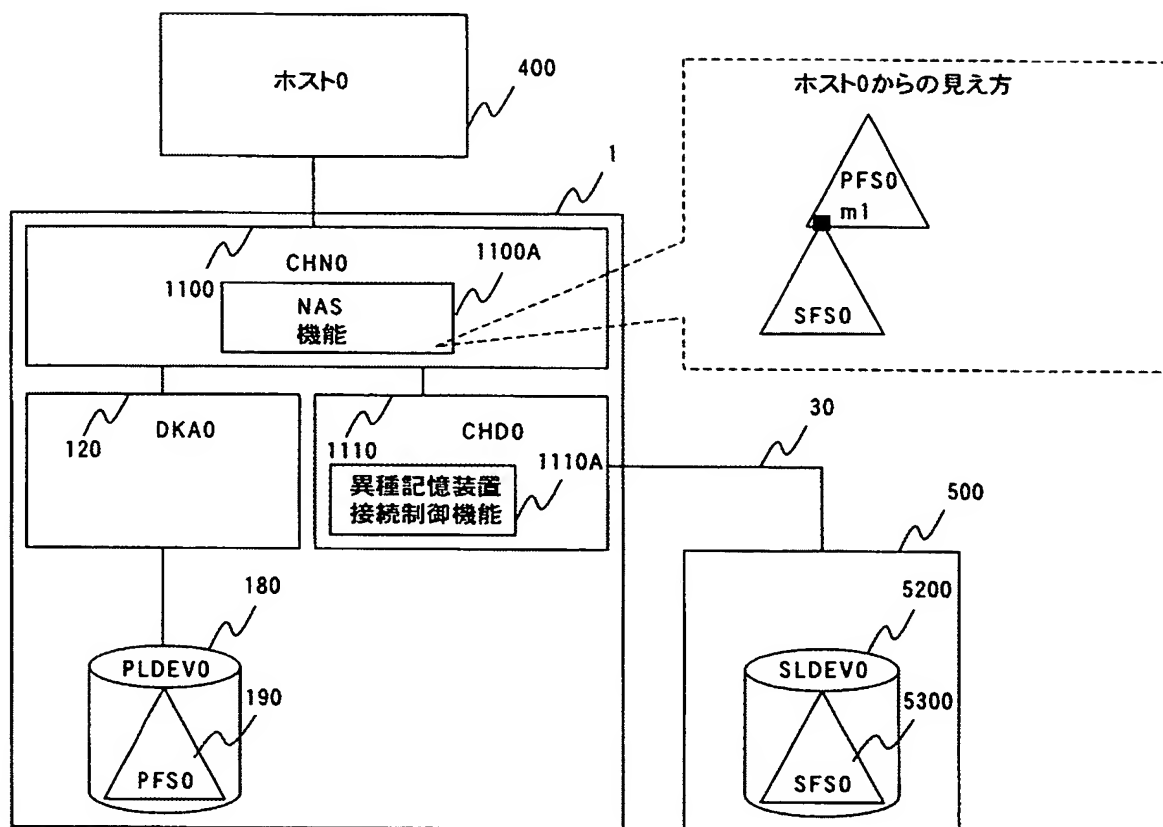
【図 12】

図12



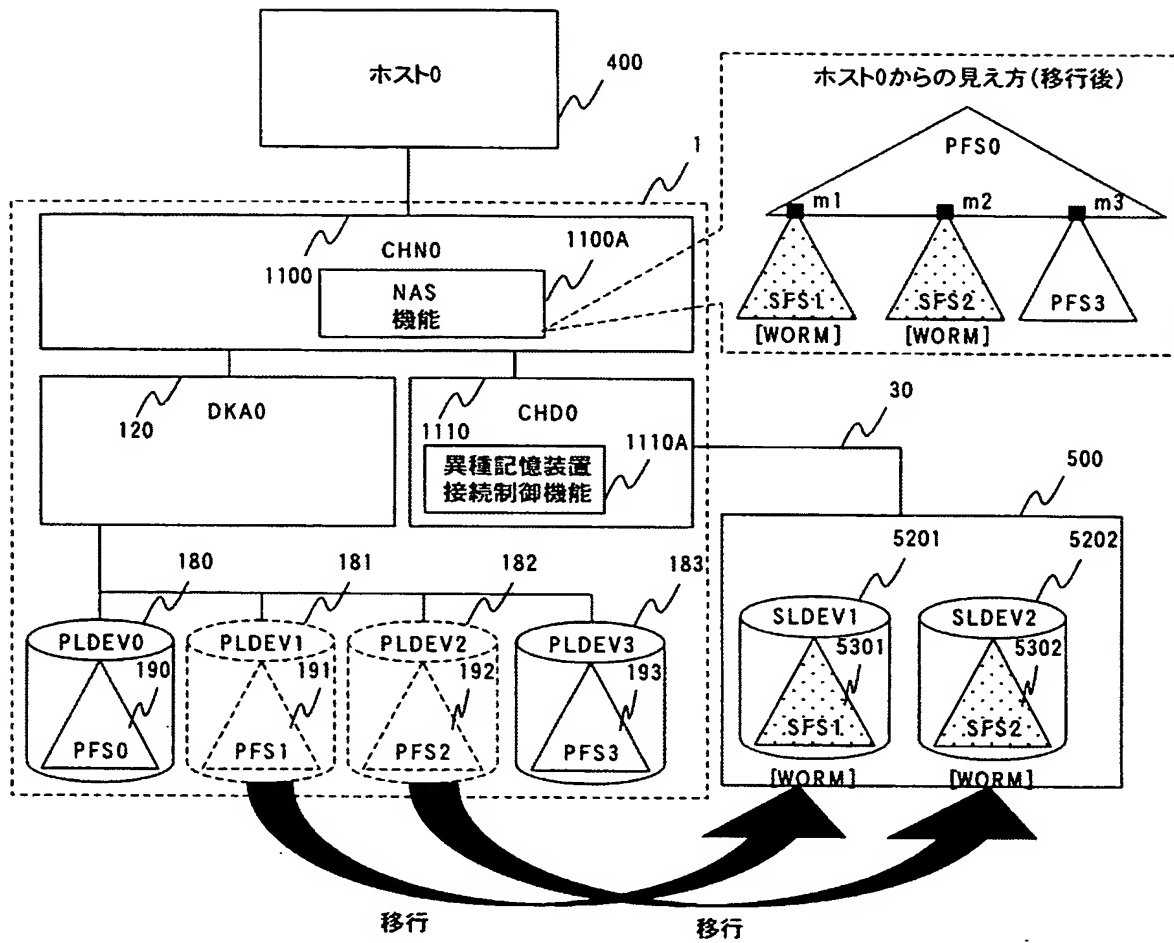
【図13】

図13



【図 14】

図14



【図 15】

図15

(a)2003年1月時点

LV	LUN	Storage	LDEV	WORM	Migration	FS	Remarks
0	0	STR0	PLDEV0	0	0	PFS0	home
1	1	STR0	PLDEV1	0	0	PFS1	2003 January

ボリューム管理テーブル

(b)2003年2月時点

LV	LUN	Storage	LDEV	WORM	Migration	FS	Remarks
0	0	STR0	PLDEV0	0	0	PFS0	home
1	1	STR1	SLDEV1	1	1	SFS1	2003 January
2	2	STR0	PLDEV2	0	0	PFS2	2003 February

ボリューム管理テーブル

(c)2003年3月時点

LV	LUN	Storage	LDEV	WORM	Migration	FS	Remarks
0	0	STR0	PLDEV0	0	0	PFS0	home
1	1	STR1	SLDEV1	1	1	SFS1	2003 January
2	2	STR1	SLDEV2	1	1	SFS2	2003 February
3	3	STR0	PLDEV3	0	0	PFS3	2003 March

ボリューム管理テーブル

【書類名】 要約書

【要約】

【課題】

大量のファイルを集約管理することができるが、容量スケーラビリティは接続できる磁気ディスク装置の台数や磁気テープ装置の台数により制限され、時間と共に増加していく大量の情報を長期に渡り十分に管理することができない。

【解決手段】

記憶装置システムは、異種記憶装置を接続し制御できるよう構成されており、記憶装置システムの内部に備える記憶領域と、異種記憶装置が備える記憶領域に対してファイルシステムを構築する。

【選択図】 図 1 3

認定・付加情報

特許出願の番号	特願 2004-037596
受付番号	50400240388
書類名	特許願
担当官	第七担当上席 0096
作成日	平成16年 2月17日

<認定情報・付加情報>

【提出日】	平成16年 2月16日
-------	-------------

出 願 人 履 歴 情 報

識別番号 [000005108]

1. 変更年月日	1990年 8月31日
[変更理由]	新規登録
住 所	東京都千代田区神田駿河台4丁目6番地
氏 名	株式会社日立製作所